Identification of the period of stability in a balance test after stepping up using a simplified cumulative sum

Doha Safieddine[1], Aly Chkeir[1], Cyrille Herlem[1], Delphine Bera[2], Michèle Collart[3], Jean-Luc Novella[4], Moustapha Dramé[4], David J. Hewson[1,5], Jacques Duchêne[1]

1. University of Technology of Troyes (UTT), Troyes, France.
2. Les Arcades Prevention Centre, Troyes, France
3. Troyes Hospital Centre, Troyes, France
4. Reims University Hospital, Reims, France
5. University of Bedfordshire, Luton, United Kingdom

Keywords: Fall risk, stabilisation, balance, change detection, cumulative sum

Address for correspondence:

Professor David Hewson

University of Bedfordshire

University Square

Luton LU1 3JU

Bedfordshire

United Kingdom

Tel: 0044 7525 616645

Email: david.hewson@beds.ac.uk

Abstract

Falls are a major cause of death in older people. One method used to predict falls is analysis of Centre of Pressure (CoP) displacement, which provides a measure of balance quality. The Balance Quality Tester (BQT) is a device based on a commercial bathroom scale that calculates instantaneous values of vertical ground reaction force (Fz) as well as the CoP in both anteroposterior (AP) and mediolateral (ML) directions. The entire testing process needs to take no longer than 12s to ensure subject compliance, making it vital that calculations related to balance are only calculated for the period when the subject is static. In the present study, a method is presented to detect the stabilization period after a subject has stepped onto the BQT. Four different phases of the test are identified (stepping-on, stabilization, balancing, stepping-off), ensuring that subjects are static when parameters from the balancing phase are calculated. The method, based on a simplified cumulative sum (CUSUM) algorithm, could detect the change between unstable and stable stance. The time taken to stabilise significantly affected the static balance variables of surface area and trajectory velocity, and was also related to Timed-up-and-Go performance. Such a finding suggests that the time to stabilise could be a worthwhile parameter to explore as a potential indicator of balance problems and fall risk in older people.

1. Introduction

The European population has been predicted to age over the coming decades, with an estimated 30% of the population aged over 65 years by 2050. The increase for people aged over 80 years for the same time period is even more rapid, rising to almost 11% of the European population [1]. Although such figures are positive in terms of a greater life expectancy, living longer also increases the risk of many diseases, which in turn leads to increased healthcare costs for the older population. In addition to the increased incidence of chronic pathologies associated with ageing, older people are also at risk of falls [2-6]. The number of fallers has been estimated at 30% for people aged over 65 years, rising to 50% for those aged over 80 years [7]. The large number of fallers has an impact on medical costs as well as a social impact, due to the increased frailty and vulnerability of fallers.

On the positive side, multi-factorial intervention programs have been shown to reduce fall risk [8]. However, such programs can only be implemented in a cost-effective manner if the most at-risk individuals can be identified. This could be achieved by detecting specific fall-risk factors, such as muscle weakness, gait impairment, or underlying balance problems.

In respect to the clinical evaluation of balance, a number of tests are frequently used, such as the Tinetti test [5, 9], the Timed Get Up and Go (TUG) test [10], the Berg Balance Scale [11, 12], and the One-Leg Stance (OLS) test [13]. Although these tests are easy to implement and do not require high-level technology, they are more suited to use in a clinical setting, as they generally require the presence of a trained evaluator. Both the TUG and OLS could be used in remote monitoring, however an additional person would be required to start and stop the timing device used. Such a restriction makes these tests expensive to perform regularly, while the time between successive examinations might be too long to detect any degradation before a fall occurs. Another drawback of most of those clinical tests is that they produce an indirect measure of balance quality, or sometimes even a binary result of “at risk” or “not at risk” (see for instance the “Stop Walking When Talking” test [14]). These tests also have a ceiling effect, with many subjects producing the maximum score when tested, making them less discriminatory than other tests.

One alternative to clinical tests of balance is to use laboratory-based measures of balance such as those that use force plates to assess balance quality based on the displacement of the Centre of Pressure (CoP) [15]. Such tests produce a multitude of parameters that are related to underlying pathologies, and can also identify balance problems [16-18]. Many devices are available to assess postural sway, including commercial force plates that automatically process CoP data and produce a wide range of related parameters. However, force plates are too expensive to be implemented as part of a large-scale prevention protocol. One solution is the Balance Quality Tester (BQT), which was developed as a low-cost balance assessment tool based on a commercial bathroom scale [19]. The BQT provides instantaneous measurement of vertical ground reaction force (Fz), and can estimate the position of the CoP (SBP) in both anteroposterior (AP) and mediolateral (ML) directions. The BQT was compared to a force plate, with high levels of validity and agreement observed between the two devices [20]. The absolute differences in the mean values of the two parameters studied for the BQT and a force plate were 0.28% for the surface area of CoP displacement and 0.62% for the mean velocity of the CoP displacement trajectory, with lower values for the BQT. Both devices had similar test-retest reliability, as assessed using interclass correlation coefficients (ICC), with 0.80 and 0.83 for the BQT and 0.80 and 0.73 for the force plate, for surface area and trajectory velocity, respectively.

The BQT has wireless communication capability, and also provides users with their weight, which is one the Fried criteria for physical frailty [21]. In the first version of the BQT, five variables were computed [19]. These were the delay before stepping onto the BQT, the velocity of weight transfer onto the BQT, the surface area of CoP displacement, the mean velocity of the CoP displacement trajectory, and the coefficient of variation (CV) of the Fz signal during the stabilization phase, which was arbitrarily defined as a 2-second period after stepping onto the device.

An empirical score of balance was then deduced using four of these five variables, depending on the experimental protocol. In the absence of an evaluator, the first four variables were used, however, if an evaluator was present, the delay before stepping-on was replaced by the CV, as the delay is heavily influenced by the evaluator's presence.

One of the advantages of the BQT is the relatively short time in which the static balance period (SBP) is calculated. This 12-second duration was chosen after a usage analysis in the ActivAgeing Living Lab (unpublished data). Testing durations longer than 12 seconds decreased the adherence of subjects with respect to daily use of the device. In addition, an SBP of 10 seconds was shown to be more reliable than shorter durations of 5 and 2.5 seconds [22]. The 12-second duration included the time taken to stabilize once on the device (2 seconds), followed by 10 seconds for static balance assessment. However, although such a short testing protocol is preferable in terms of subject compliance, it is possible that some subjects might not be stable at the end of the 2-second period before the SBP is calculated. It is necessary, therefore, to determine the exact period when subjects are stable in order in be certain that CoP parameters are relevant. In addition, the time taken to reach stabilization could be used as an additional variable related to balance quality.

The aim of the present paper is to evaluate the accuracy of a new method to estimate the duration of the stabilization phase in a balance test, with the cut-off point then used to select the period when static balance can be analysed. The hypothesis to be tested is that the CUSUM method will provide a more accurate estimation of the SBP than the use of an arbitrary 2-second window. A comparison will be made between the parameters calculated using the new segmentation and those obtained using the empirical segmentation from previous studies [19, 20, 23].

2. Methodology

2.1 Subjects and protocol

One hundred and seventy-five older subjects were recruited, with their characteristics shown in Table I. Ethical approval was obtained from the regional ethics committee for biomedical research, and all subjects gave written informed consent (CPPRB 2019-A00146-39). No subjects reported any musculoskeletal or neurological symptoms that would have prevented them from participating in the study. Each subject was evaluated once using a similar protocol to that used in previous studies [19]. The only difference was that an improved version of the BQT was used, with a sample frequency of 100Hz rather than the 16Hz frequency of the previous model. The new version also provided the exact time when a subject's weight was displayed on the BQT screen.

2.2 Data acquisition

The device is based on a commercial bathroom scale that is equipped with four pressure sensors that provide ground reaction force at each of their locations (Fig.1) (see [19] for more details). The scale was adapted to provide access to the raw data produced by the sensors, and to operate using a simple experimental protocol designed for self-measurement at home. Subjects started by standing in front of the BQT to activate an infrared (IR) sensor that detected the presence of the person. Activation of the IR sensor turned the BQT on and "0.0" was displayed, which was the signal for the subject to step onto the BQT. The subject then remained standing as still as possible until body weight was displayed, which took 12s. Once their weight had been displayed, subjects stepped off the BQT.

The instantaneous position of the CoP was defined as the barycentre of the four vertical ground reaction forces measured by the sensors, which were not filtered. Measurements of displacement AP and ML were also calculated using the same method. Vertical ground reaction force Fz was taken as the sum of the four ground reaction forces measured by the individual sensors (Fig.1).

2.3 Signal segmentation

In a previous version of the device, the exact time when weight was displayed was not provided. This meant that it was difficult to determine when the SBP ended and the preparation for stepping off began. The end of the SBP was estimated as occurring two seconds before the moment when Fz dropped below 90% of a subject's body weight. The same method was used to estimate when the SBP started. This estimation could have resulted in a loss of part of the relevant signal, something that could be critical for a total recording time of 12s. Furthermore, it could have resulted in the inclusion of part of the stabilization phase within the SBP, thus inducing a bias in the estimation of the SBP variables. Finally, the variable used to reflect the stabilization phase itself, the coefficient of variation, was only calculated for the first two seconds after Fz exceeded 90% of body weight during the stepping up phase.

Given that the new version of the BQT provides the time when the weight is displayed, it is possible to define a new segmentation algorithm based on three assumptions (Fig. 2). Firstly, it was assumed that the stabilization phase starts when Fz exceeds the person's weight for the first time (t_p). Secondly, that the SBP ends when the person's weight is displayed (t_w), and finally that the statistical features of the stabilization phase will differ from those of the SBP. Based on the third assumption, the starting point of the SBP was defined as the time when there is a statistical change in Fz.

Two methods could be used to detect a statistical change in Fz. Firstly, a different distribution function could be detected before and after the change time, or secondly, a modification in the value of a parameter of the distribution, such as the variance could be detected. One algorithm that has been shown to detect such statistical changes is the Cumulative Sum (CUSUM) (see [24] for a review of different methods). The CUSUM algorithm is based on the following hypothesis:

Let $X(x_1, x_2, \ldots, x_n)$ be a time series that includes a possible change in the vector of parameters θ of its probability density function f_θ at a time $\boldsymbol{k}$. The hypothesis test is then:

$$H_0 : \theta = \theta_0, \; for \; 1 \le i \le n \tag{1}$$

against

$$H_1 : \begin{cases} \theta = \theta_0, \; for \; 1 \le i \le k \\ \theta = \theta_1, \; for \; k+1 \le i \le n \end{cases} \tag{2}$$

The main detection algorithms are based on the sum of the logarithm of the likelihood ratio, which has been demonstrated as an appropriate statistic in detection theory [24]:

$$S^j(x_1, x_2, \ldots, x_j) = \sum_{i=1}^{j} \ln \frac{f_{\theta_1}(x_i / x_{i-1}, \ldots, x_1)}{f_{\theta_0}(x_i / x_{i-1}, \ldots, x_1)} \tag{3}$$

In the case of an independent Gaussian sequence with zero mean and a change in variance, each term of S can be expressed as:

$$s_i = \frac{1}{2}\left[\ln\frac{\sigma_0^2}{\sigma_1^2} + x_i^2\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)\right] \quad (4)$$

In the present work, it is known that there should be a change in the signal between the stabilization phase and the SBP, meaning the problem is twofold. Firstly, is it possible to detect the change, given that the parameters of the probability density function are not *a priori* known and have to be estimated? Secondly, if the change is detectable, can the time position t_k of the change be computed?

When both feet are steady on the scale, it was assumed that the only parameter to change would be the variance, σ^2. Given that σ_1 and σ_0 are unknown, both parameters need to be estimated at the beginning (stabilization) and at the end (SBP) of the sequence, respectively. The estimation of the variance σ_0 at the end of the SBP just before weight is displayed can be considered as valid due to the stationarity of this zone. However, the estimate σ_1 at the beginning of the time series is debatable as the time to become stable could be very short, while the signal is unlikely to be stationary. Nevertheless, despite these limitations, σ_1 and σ_0 were estimated for the 2-s intervals following t_p and prior to t_w, respectively (Fig. 2, lower trace). An argument for the validity of such an approach is provided in the discussion section.

Another difficulty is that successive samples are probably not independent given the sampling frequency of 100Hz (N=200), thus leading to an erroneous confidence interval. A simple empirical method to produce an independent time series is to under-sample the original series, with the rate of under-sampling defined from the autocorrelation function. Following the hypotheses stated above, the first step was to verify that estimates of σ_1 and σ_0 agreed with the hypothesis $\sigma_1 > \sigma_0$. This hypothesis was tested using Fisher's distribution:

$$\sigma_1 > \sigma_0 : \left(\frac{\hat{\sigma}_1^2}{\hat{\sigma}_0^2}\right) > F_{n,n}\Big|_{p=0.95} \quad (5)$$

If this test holds true, it follows that a change occurs at some point within the time series, hence reducing the CUSUM problem to the determination of the change time t_k. The change time t_k is detected based on the behaviour of the log-likelihood ratio S^j, which has a negative slope before the change time t_k and a positive slope thereafter. Accordingly, the detection function at a time *j* was calculated as:

$$g_j = S^j - \min_{1 \le i \le j} S^i \qquad (6)$$

The CUSUM algorithm was applied in the reverse direction, starting from the moment that weight was displayed, then going back to the starting point of the stabilization phase (Fig. 2). This method was possible, as change detection was made off-line, while it was assumed that the SBP phase was stationary. Finally, it was assumed that there would be only one change in the time series, resulting in the following computation for the detection time $\boldsymbol{t_k}$:

$$t_k = \max\left\{t_j : g_j = 0\right\}$$

2.4 *Extraction of variables*

Two variables were computed from the Fz signal, namely the rise rate (RR) and the duration of the stabilization segment (ZD). With respect to the RR, it was defined as the average slope between 10% and 90% of bodyweight during the stepping up phase (Fig. 3). This variable takes account of any hesitations between the contact of the first foot onto the scale and the final phase of the contact of the second foot. The presence of any inflexions or peaks in the Fz rising phase due to movements of the second foot would also increase the value of this variable. The ZD parameter was taken to be the time between t_p and t_k.

Two variables were also extracted from the CoP trajectory. The area of the SBP (SA) was estimated as the product of the standard deviation of CoP displacement in the anteroposterior and mediolateral directions multiplied by 4π, which roughly approximates an ellipse. Intuitively, this variable expresses the amount of stability during the static phase, considering both anteroposterior and mediolateral oscillations. The average velocity of the trajectory (TV) was computed as the sum of the lengths taken between successive points, divided by the SBP duration. This variable, which is equivalent to the CoP path length in a fixed time interval, has been shown to be a relevant measure of standing balance [25].

From the definitions above, three of the variables are strongly influenced by the change time (t_k) between stabilization and the SBP (SA, TV, ZD). Each of the four variables is subsequently scored empirically in the range [0-4] based on a series of thresholds (Table II). This transformation from a native value to a score has been detailed in [19] for three of the variables (RR, SA, TV). For the fourth variable, ZD, score limits were defined as the quintiles of the experimental distribution function. The validity of such an assumption is provided in the discussion section. Finally, a global empirical score was computed as a simple addition of either all four partial scores or for a subset of partial scores.

2.5 Performance evaluation

The hypothesis behind the adoption of the CUSUM method is that it would be possible to separate the stabilization and static balance periods of the signal, something that is necessary to use traditional stabilogram parameters. The variance in each part of the signal was used to compare the performance of the CUSUM method with the original and arbitrary use of a two-second cut-off. With respect to the CUSUM method, the stabilization phase (S_z) was taken as the period between t_p and t_k, while the SBP (S_b) of the signal was taken as the period between t_k and t_w. For the original method, S_z was taken to occur between t_p and t=2s, while the S_m was taken to occur between t=2s and t_w.

The variance for each of these periods was calculated, as well as the ratio Q between the two periods, where:

$$Q = \frac{variance\ S_z}{variance\ S_m}$$

If the CUSUM method detected a shorter stabilization than 2s, Q_{Tk} would be less than Q_{2s}. In contrast, when the CUSUM detected a stabilization longer than 2s, Q_{Tk} would be greater than $Q_{2s.}$

The SA and TV variables were also compared with and without optimization using the CUSUM algorithm. The performance of the CUSUM method was also compared with the results of the OLS test for balance and the TUG test for mobility. The stabilisation durations of the subjects were used to classify the subjects into four groups, with each group containing one quartile of the subjects with respect to their stabilisation duration (fastest 25%, second fastest 25%, third fastest 25%, and the slowest 25%).

Normality was verified using the Lilliefors test [26], given that the application of the more commonly-used Kolmogorov-Smirnov test for normality assessment is debatable when the distribution parameters have to be estimated from sample data. Data was displayed as mean ± SD for normal data and as median (16-84 percentile range) for non-normal data. A comparison between methods was made using either the Wilcoxon test for non-normal data, or a t-test for normally distributed data, depending on the result of the Lilliefors test. The comparison between the stabilisation duration and the TUG was performed using Analysis of Variance (ANOVA) with the bias-corrected and accelerated (BCa) bootstrap method to produce unbiased estimates of the confidence limits around the mean [27]. The significance level was set to $p<0.05$ for all tests.

3. Results

3.1 Preliminary tests

The rate of under-sampling was estimated using the auto correlation length in all 175 subjects. Based on these results, an under-sampling rate of seven was used for all subsequent analyses. After under-sampling, Fisher's test was used to compare the variances of the 2-s segments at the beginning (S_1: stabilization) and the end (S_0: SBP) of Fz after segmentation (Fig. 2, lower trace).

Significant differences were observed for each of the 175 signals. The Lilliefors test was then used on the same segments, S_1 and S_0, to determine normality. Most of the S_0 segments were normally distributed (90.9%), whereas almost none of the S_1 segments were normally distributed (1.1%). A chi-square test showed a significant difference between the two segments with respect to the number of normal distributions (p=0.000).

3.2 Signal segmentation

Two typical detections of the change time are shown in Fig.4 (upper tracings). The Fz signals are plotted backwards from t_w to t_p, which replicates the method used by the CUSUM algorithm. The behaviour of the detection function g is also shown (Fig.4: lower tracings). Following the assumption that there is only one abrupt change in each time series, change time t_k was defined as the last time when g crosses zero, thus avoiding the use of a threshold. The frequency distribution of t_k for all subjects, which is shown in Fig.5, has a median value of 1.86s. This median was not found to differ significantly from the arbitrary 2s stabilization duration used previously (Wilcoxon signed-rank test: Z=-.022, p=0.999). However, it can be seen from the histogram that many subjects had stabilisation durations that differed markedly from the arbitrary 2s cut off chosen previously. In total, 65.1% of subjects had a stabilisation duration that differed from the 2s cut off by more than 0.5 seconds.

Differences in SA and TV variables between the two methods used to calculate the stabilization period were also evaluated for differences (ΔSA and ΔTV calculated as 2s empirical segmentation - CUSUM algorithm segmentation). Normality was verified using the Lilliefors test (n=175, Dcrit=0.067, p=0.05). Neither of the two distributions were normally distributed (Dmax=0.078 and 0.102 for ∆SA and ∆TV, respectively). Accordingly, the Wilcoxon signed-rank test was used for both ∆SA and ∆TV with neither variable differing significantly from zero (z=-.002, p=0.999 and z=0.1, p=0.920 for ∆SA and ∆TV, respectively).

3.3 Performance evaluation

The overwhelming majority of subjects had change times that differed markedly from the 2-sec value previously used in the empirical method. With respect to the absolute magnitude of the differences between the change time and 2 sec, the median absolute difference was 0.72 sec (16-84 percentile range 0.17-1.47 sec).

This result is reflected in the median change time observed for the CUSUM method of 1.86 sec (16-84% range: 1.04-3.26 sec). The ratio *Q* between the stabilisation and static balance zones was significantly different for the two methods using the Wilcoxon signed rank test, with a median value of 74.5 for Q_{TK} compared to 53.5 for Q_{2S} (p=0.000).

The differences in the amount of variance in each of the periods effected the scores calculated for balance quality, based on the variables shown in Table II. In the previous method, the median score was 10 (16-84% range: 7-13), whereas in the CUSUM method, the median score was 9 (16-84% range: 6-12). This difference was significantly different from zero using the Wilcoxon signed rank test (p=0.000).

The results of the TUG performance for the four quartiles of stabilisation duration are shown in Table III). There was a significant effect of the stabilisation duration with respect to the TUG performance (F=5.318, p=0.002). The performance of the slowest group was significantly worse than for all other groups, however there were no differences between the three fastest groups with respect to TUG performance.

The effect of stabilisation duration was also evaluated for the two SBP variables, SA and TV (Table IV). There was a significant effect of the stabilisation duration with respect to SA (F=5.559, p=0.001), but not for TV (F=1.662, p=0.177). The fastest group to stabilise according to the CUSUM method had significantly greater values for SA than did the two slowest groups. There were no other significant differences between groups for SA.

4. Discussion

The primary use of the BQT is for home-based assessment of balance quality. In previous work, the BQT was as effective as standard balance tests in discriminating between community-dwelling older people and nursing home residents [23]. To improve usability a 12-s testing duration was chosen, after discussions with potential users. The usability of a short test duration was born out in a previous study in which 22 older subjects were tested on a near-daily basis for at least 12 months [19]. In addition, despite the short-testing duration, parameters were shown to have high reliability, with intra-class correlation coefficients exceeding 0.80 for the parameters tested [20]. One potential drawback of the short testing duration is that subjects might not be stable at the end of 2s period previously used to detect the start of the static phase of balance. This study aimed to determine when subjects became stable so that only static balance was used to calculate SBP parameters.

When the CUSUM method to was used to detect stability, there was a marked difference in the time taken to stabilise on the device compared to the arbitrary 2s duration used previously. Close to two-thirds of subjects had stabilisation durations that differed from 2s by more than 0.5s. Indeed, many subjects took so long to stabilise, it is doubtful whether any variables extracted from the SBP of the signal would be worthwhile. It seems clear that an accurate detection of the SBP is crucial for the BQT to be used to evaluate balance. It seems logical that those subjects that took longer to stabilise on the device, might have other balance-related problems and could therefore be worthy of further investigation. The time taken to stabilise could be a worthwhile parameter to explore as a potential indicator of balance problems and fall risk in older people. There were also several methodological issues related to the successive steps of this algorithm that need to be addressed.

4.1 Methodological considerations

One of the hypotheses behind the CUSUM algorithm is the normality of the analysed signals [24]. Normality was well established for the SBP, but not for the stabilization phase. Such a result could lead to degradation in the CUSUM algorithm performance as CUSUM performance for non-normal distributions remains an open question [28]. However, it has also been shown that CUSUM procedures are less sensitive to non-normality than other methods [28, 29], which might explain the satisfying results produced by the CUSUM algorithm in the present study.

The same limitation applies when estimating σ_0 and σ_1. Although the estimation of σ_0 at the end of the SBP can be considered as valid (stationary and normally distributed time series), it is not the case for σ_1 in the 2s window at the beginning of the stabilization phase. The stabilization phase is often much shorter, with a median duration of 1.7 and cannot reasonably be considered as stationary. Despite these limitations, the 2s window was used to estimate σ_1. In the case where the stabilization phase exceeds 2s, it could be hypothesized that the time series used for σ estimation would be more likely to be stationary. In contrast, if stabilization occurs quicker than 2s, σ_1 would be underestimated but would still respect the condition $\sigma_1 > \sigma_0$ required for Fisher's test.

When the results of the study are considered, it can be observed that the median of the sample histogram of 1.86s did not differ significantly from the 2s segment duration, which was the stabilization duration empirically defined in previous studies [19]. However, the histogram was skewed towards higher values, indicating that in many cases part of the stabilization phase was included in the SBP, thus creating a potential bias in the estimation of the SBP variables SA and TV.

4.2 Empirical score

The empirical method used to deduce partial scores from the RR, SA and TV variables [19] was not evaluated in the present study as the aim was to identify any improvement provided by an optimal Fz segmentation. In respect to ZD, the thresholds for the partial scores of 0-4 were empirically defined as the quintiles of the ZD experimental distribution, which was also the method used to derive the partial scores for the other variables that make up the balance score. It seems clear that optimizing the segmentation of the raw variables into partial scores should be the aim of future work, although it should be kept in mind the fact that there is no "Gold Standard" as a reference for the optimization process. The same argument could be applied for the combination of the partial scores to produce the global empirical score.

The results observed for the two SBP variables, SA and TV, with respect to stabilisation duration, imply that the method to quantify balance should be readdressed due to the use of the CUSUM method. In long duration assessments of static balance, higher values of both SA and TV are associated with worse balance, whereas in the present study, greater values of SA were observed for those subjects who could stabilise themselves more quickly. This finding could be due to the short recording duration used, as noted in previous work [22]. It seems clear that the balance score will need to be readdressed considering these results. It would also be of interest to examine other characteristics of the different phases of the protocol, especially the SBP, by taking into account, for instance, the possible nonlinear properties of the signal [22] or its dynamic behaviour [30]. However, the extraction of variables that define nonlinear characteristics would be limited by the short duration of the time series available.

The findings of the present study that the time to stabilisation is strongly related to TUG performance offers many interesting perspectives. It would be worthwhile determining whether the ability to stabilise after a step could predict fall risk. It should also be noted that balance quality is only one factor among many in relation to fall risk. Many other clinical tests related to fall risk, such as gait velocity [21] or the Timed Up and Go test [11], could also be used. For instance, in a recent study, it was shown that walking speed correlated well with frailty, which in turn is related to fall risk [31]. Further study is needed in which the time to stabilise is compared with other tests of fall risk, preferably in longitudinal studies.

The present study has some limitations. Firstly, the effectiveness of the CUSUM method in detecting the SBP was only compared to the 2-s stabilization period used previously. It is possible that other methods might also have produced acceptable results, like those of the CUSUM. Secondly, the empirical scoring method was based on quintiles from the experimental distribution in the present study. Results could have differed if a different population was used for the experiment. Finally, none of the four parameters included in the balance score address potential issues of the non-linearity of the balance, meaning that additional work is required in this area.

5. Conclusion

In the present study, a segmentation algorithm was applied to vertical ground reaction forces obtained from a modified bathroom scale. The algorithm produces segments corresponding to different phases of the weighing protocol (stepping up, stabilization, SBP). The effect of this segmentation was investigated in respect to the relevance of the extracted variables when compared to an empirical segmentation used previously. The new algorithm could identify when subjects were able to stabilise themselves on the bathroom scale, with many subjects being unable to do so within the 12s currently used by the device for its testing protocol. The time taken to stabilise changed the way in which the balance score needs to be calculated, and was associated with TUG performance. The parameter could be worthwhile exploring as a potential indicator of balance problems and fall risk in older people.

6. Acknowledgments

This work was supported by the Champagne-Ardenne Regional Council (CRCA) and the European Regional Development Fund under the Collaborative Research Program (Domo-Grip Project, grant reference numbers CRCA E201012437 and FEDER E201013375). The authors would like to thank all clinicians, clinical research associates and technicians for their involvement in the device design and the protocol definition, set-up and monitoring, especially the three clinical sites (Reims University Hospital, Troyes Hospital Centre and the Prevention Centre “Les Arcades”).

7. References

[1] European Commission: Directorate-General for Economic and Financial Affairs. The 2005 EPC projections of age-related expenditure (2004–50) for the EU-25 Member States. European Economy: European Commission: Directorate-General for Economic and Financial Affairs, Special Report No 4; 2005.

[2] Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls: a prospective study. JAMA. 1989;261:2663-8.

[3] Raîche M, Hébert R, Prince F, Corriveau H. Screening older adults at risk of falling with the Tinetti balance scale. The Lancet. 2000;356:1001-2.

[4] Rubenstein LZ, Josephson KR. The epidemiology of falls and syncope. Clinics in Geriatric Medicine. 2002;18:141-58.

[5] Tinetti ME, Speechley M, Ginter SF. Risk Factors for Falls among Elderly Persons Living in the Community. New England Journal of Medicine. 1988;319:1701-7.

[6] Wood BH, Bilclough JA, Bowron A, Walker RW. Incidence and prediction of falls in Parkinson's disease: a prospective multidisciplinary study. Journal of Neurology, Neurosurgery & Psychiatry. 2002;72:721-5.

[7] World Health Organization. Global health and ageing. World Health Organization; 2011.

[8] Tinetti ME, Baker DI, McAvay G, Claus EB, Garrett P, Gottschalk M, et al. A multifactorial intervention to reduce the risk of falling among elderly people living in the community. New England Journal of Medicine. 1994;331:821-7.

[9] Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. J Am Geriatr Soc. 1986;34:119-26.

[10] Mathias S, Nayak U, Isaacs B. Balance in elderly patients: the "get-up and go" test. Arch Phys Med Rehabil. 1986;67:387-9.

[11] Shumway-Cook A, Baldwin M, Polissar NL, Gruber W. Predicting the probability for falls in community-dwelling older adults. Physical therapy. 1997;77:812-9.

[12] Tsang WW, Hui-Chan CW. Effects of Tai Chi on joint proprioception and stability limits in elderly subjects. Med Sci Sports Exerc. 2003;35:1962-71.

[13] Vellas BJ, Wayne SJ, Romero L, Baumgartner RN, Rubenstein LZ, Garry PJ. One-leg balance is an important predictor of injurious falls in older persons. J Am Geriatr Soc. 1997;45:735-8.

[14] Hyndman D, Ashburn A. “Stops walking when talking” as a predictor of falls in people with stroke living in the community. Journal of Neurology, Neurosurgery & Psychiatry. 2004;75:994-7.

[15] Haas BM, Burden AM. Validity of weight distribution and sway measurements of the Balance Performance Monitor. Physiotherapy Res Int. 2000;5:19-32.

[16] Tsang WW, Hui-Chan CW. Sensorimotor control of balance: a Tai Chi solution for balance disorders in older subjects. Medicine and sport science. 2008;52:104-14.

[17] Maki BE, Holliday PJ, Topper AK. A Prospective Study of Postural Balance and Risk of Falling in An Ambulatory and Independent Elderly Population. Journal of Gerontology. 1994;49:M72-M84.

[18] Thapa PB, Gideon P, Brockman KG, Fought RL, Ray WA. Clinical and Biomechanical Measures of Balance Fall Predictors in Ambulatory Nursing Home Residents. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 1996;51A:M239-M46.

[19] Duchene J, Hewson DJ. Longitudinal evaluation of balance quality using a modified bathroom scale: usability and acceptability. J Telemed Telecare. 2011;17:421-6.

[20] Hewson D, Duchêne J, Hogrel J-Y. Validation of balance-quality assessment using a modified bathroom scale. Physiol Meas. 2015;36:207-18.

[21] Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A-Biol. 2001;56:M146-M57.

[22] Amoud H, Abadi M, Hewson D, Michel-Pellegrino V, Doussot M, Duchêne J. Fractal time series analysis of postural stability in elderly and control subjects. J Neuroeng Rehabil. 2007;4:12.

[23] Vermeulen J, Neyens JCL, Spreeuwenberg MD, van Rossum E, Hewson DJ, Duchêne J, et al. Construct validity of a modified bathroom scale that can measure balance in elderly people. J Am Med Dir Assoc. 2012;13:665.e1-e5.

[24] Basseville M, Nikiforov IV. Detection of abrupt changes: theory and application: Prentice-Hall, Inc. Upper Saddle River, NJ, USA; 1993.

[25] Salavati M, Hadian MR, Mazaheri M, Negahban H, Ebrahimi I, Talebian S, et al. Test-retest reliability of center of pressure measures of postural stability during quiet standing in a group with musculoskeletal disorders consisting of low back pain, anterior cruciate ligament injury and functional ankle instability. Gait Posture. 2009;29:460-4.

[26] Lilliefors HW. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. J Am Stat Assoc. 1967;62:399-402.

[27] Kelley K. The effects of nonnormal distributions on confidence intervals around the standardized mean difference: bootstrap and parametric confidence intervals. Educ Psychol Meas. 2005;65:51-69.

[28] Ryan TP, Faddy BJ. The Effect of Non-Normality on the Performance of CUSUM Procedures. In: Lenz H-J, Wilrich P-T, editors. Frontiers in Statistical Quality Control 6: Physica-Verlag HD; 2001. p. 176-93.

[29] Hawkins DM, Olwell DH. Cumulative sum charts and charting for quality improvement: Springer Science & Business Media; 1998.

[30] Muir JW, Kiel DP, Hannan M, Magaziner J, Rubin CT. Dynamic parameters of balance which correlate to elderly persons with a history of falls. PLoS One. 2013;8:e70566.

[31] Theou O, Jones GR, Jakobi JM, Mitnitski A, Vandervoort AA. A comparison of the relationship of 14 performance-based measures with frailty in older women. Appl Physiol Nutr Metab. 2011;36:928-38.

8. Figure and Table Captions

Fig 1. The modified bathroom scale

Fig.2. Segmentation of the Fz signal. Upper trace: Detection of the start of stabilization (t_p) and the end of the SBP (t_w). Lower trace: Detection algorithm for the stabilization phase.

Fig.3. Calculation of the rising rate parameter.

Fig. 4. An example of two typical segmentation results: Fz plotted from t_w to t_p with the corresponding g function. The solid vertical line represents the change time. The dashed line indicates 2s from t_p [19].

Fig.5. Frequency distribution of change detection time t_k. Vertical dashed and dotted lines represent the median value and the arbitrary 2s stabilization duration, respectively.

Table I: Population characteristics

Table II: Thresholds for scoring the variables in the balance quality score

Table III: The effect of stabilisation duration on Timed-Up-and-Go performance

Table IIV: The effect of stabilisation duration on static balance parameters

Table 1

	#	Age (y)	Height (m)	Weight (kg)
Women	105	78.8 ± 5.6	1.57 ± 0.07	66.1 ± 12.7
Men	70	78.2 ± 5.2	1.71 ± 0.06	79.7 ± 11.6

Data are means ± SD

Table 2

Score value	Rising rate (RR, $kg.s^{-1}$)	Stabilization duration (ZD, s)	Stabilogram area (SA, cm^2)	Trajectory velocity (TV, $cm.s^{-1}$)
0	<60	≥2.7	≥12	≥5
1	60,<80	2, <2.7	8, <12	4, <5
2	80, <100	1.45, <2	5, <8	2.5, <4
3	100, <120	0.9, <1.45	3, <5	2, <2.5
4	≥120	<0.9	<3	<2

Table 3

	Timed-Up-and-Go (s)
Fastest 25%	10.4 (9.6 – 11.4)
Second fastest 25%	10.6 (9.5 – 11.8)
Third fastest 25%	11.0 (10.3 – 11.9)
Slowest 25%	13.5 (12.0 – 15.2)

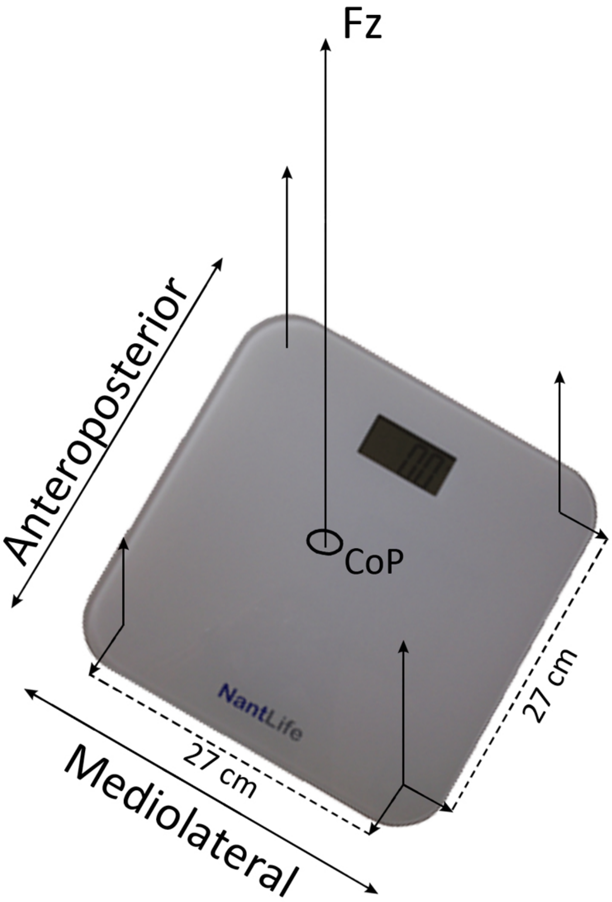
Data are bootstrapped means and 95% confidence intervals

Table 4

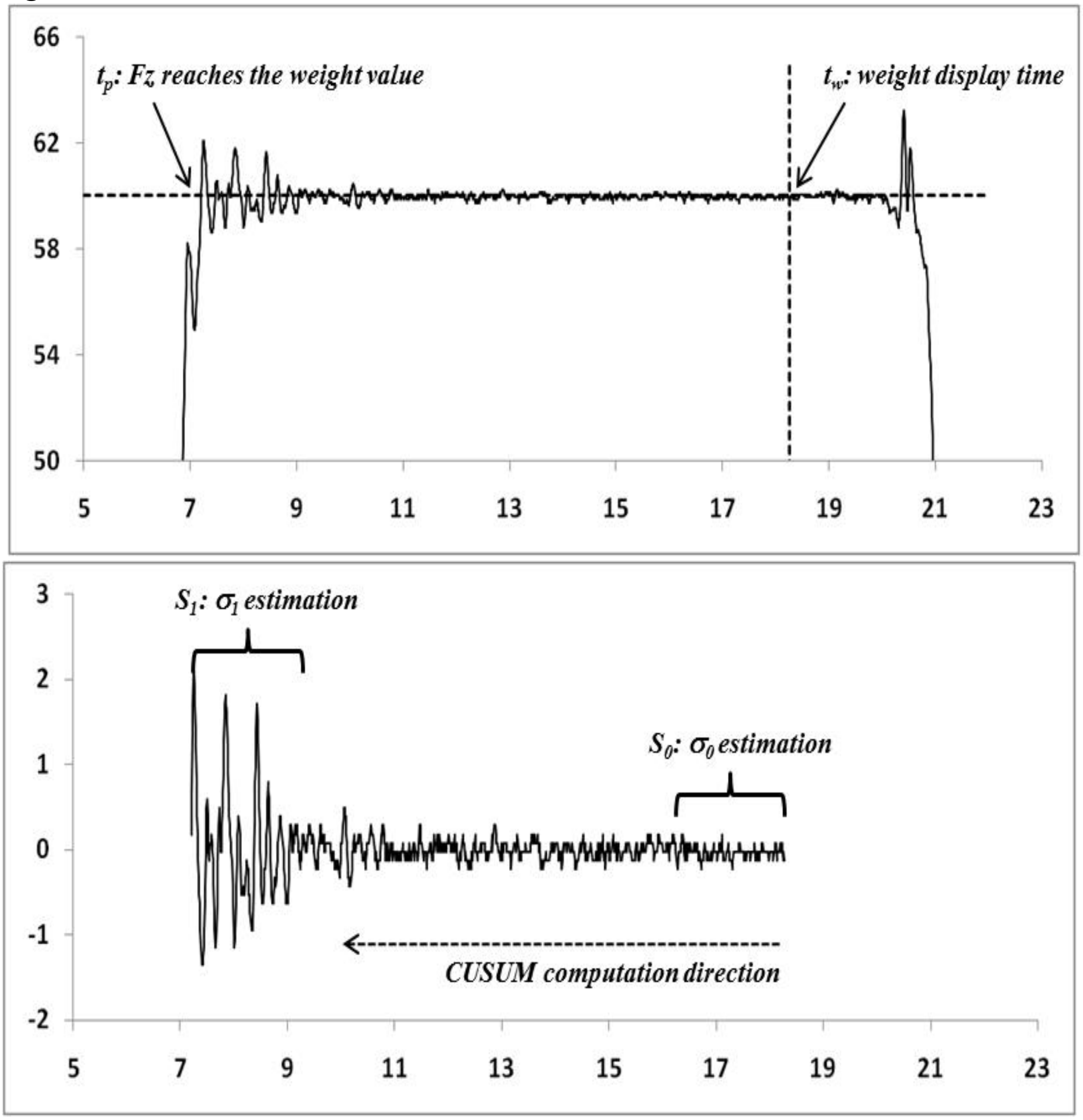
	Stabilogram Surface Area, SA (cm^2)	Trajectory Velocity, TV ($cm.s^{-1}$)
Fastest 25%	7.2 (5.7 – 9.1)	3.4 (3.0 – 3.8)
Second fastest 25%	5.8 (4.9 – 7.0)	3.0 (2.7 – 3.4)
Third fastest 25%	4.2 (3.6 – 5.0)	2.9 (2.7 – 3.2)
Slowest 25%	4.2 (3.4 – 5.0)	2.8 (2.5 – 3.3)

Data are bootstrapped means and 95% confidence intervals

Figure 1



567 **Figure 2**



568
569

Figure 3

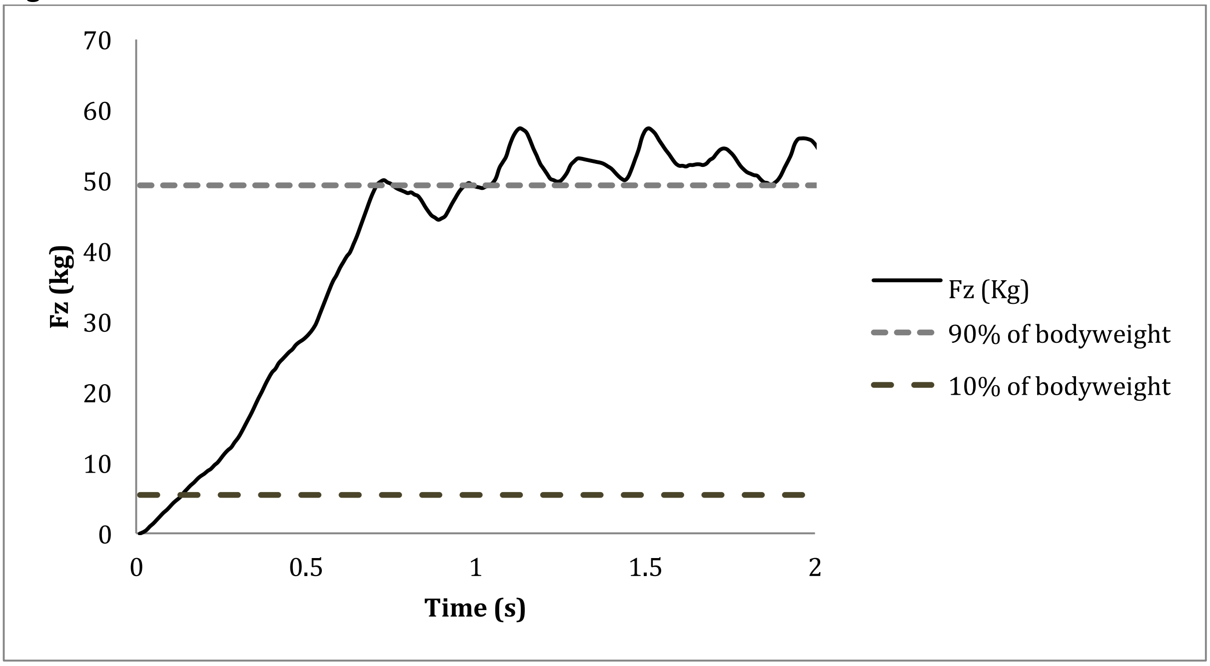
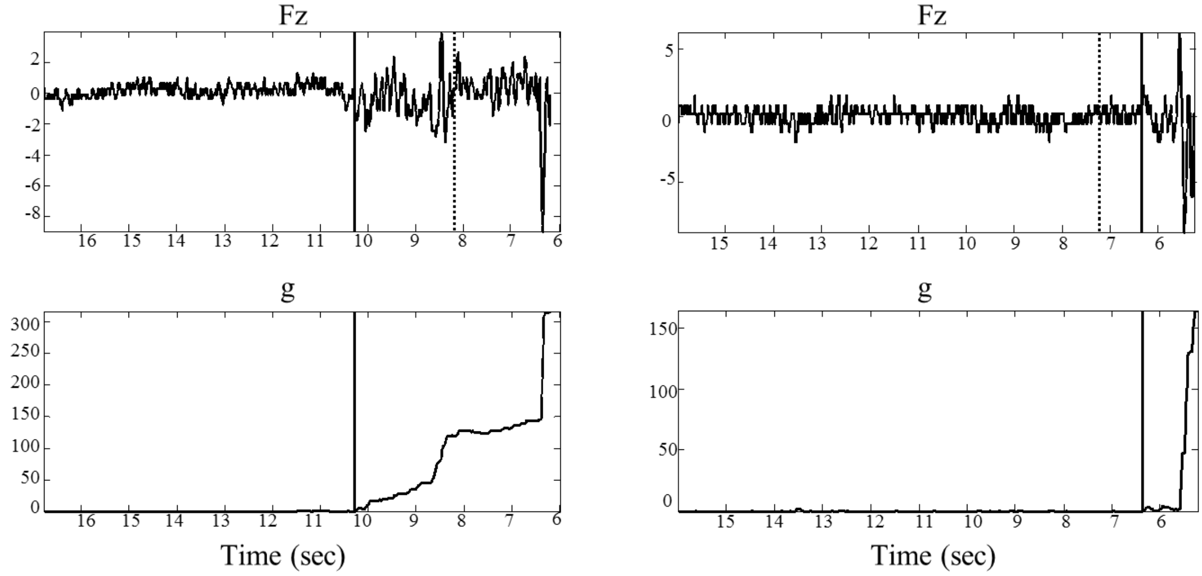
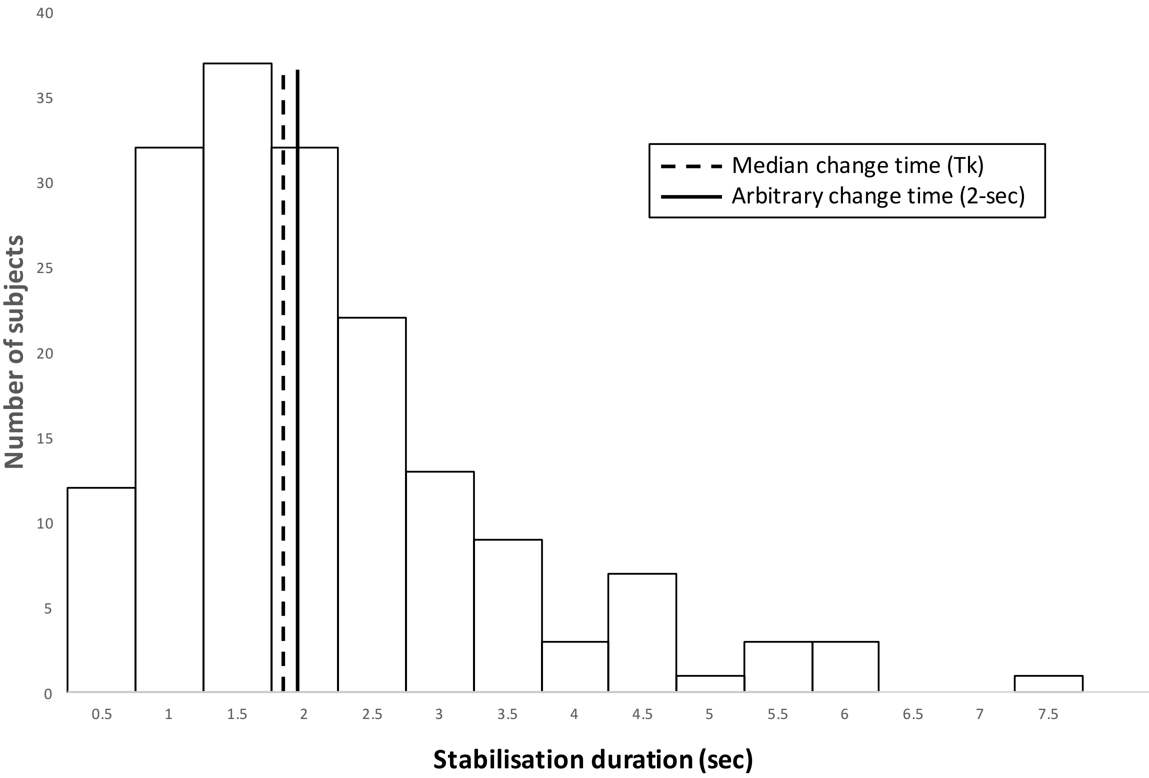

Fz (kg)
70
60
50
40
30
20
10
0
0
0.5
1
1.5
2
Time (s)
Fz (Kg)
90% of bodyweight
10% of bodyweight

Figure 4



Fz
2
0
-2
-4
-6
-8
16 15 14 13 12 11 10 9 8 7 6
g
300
250
200
150
100
50
0
16 15 14 13 12 11 10 9 8 7 6
Time (sec)
Fz
5
0
-5
15 14 13 12 11 10 9 8 7 6
g
150
100
50
0
15 14 13 12 11 10 9 8 7 6
Time (sec)

Figure 5


Median change time (Tk)
Arbitrary change time (2-sec)
Number of subjects
0
5
10
15
20
25
30
35
40
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
7.5
Stabilisation duration (sec)